DOCUMENT RESUME

ED 307 592 CS 009 674

AUTHOR McConkie, G. W.; And Others
TITLE Eye Movement Control during Reading: II. Frequency of Refixating a Word. Technical Report No. 469.
INSTITUTION Bolt, Beranek and Newman, Inc., Cambridge, Mass.; Illinois Univ., Urbana. Center for the Study of Reading.
SPONS AGENCY National Inst. of Child Health and Human Development (NIH), Bethesda, Md.; Office of Educational Research and Improvement (ED), Washington, DC.
PUB DATE May 89
CONTRACT OEG-0087-C1001
GRANT #HD18116
NOTE 28p.; For "Eye Movement Control during Reading I," see ED 286 166.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS College Students; *Eye Fixations; Higher Education; Perceptual Motor Coordination; Reader Text Relationship; *Reading Processes; Reading Research; *Word Recognition
IDENTIFIERS Saccadic Eye Movements; Word Length

ABSTRACT

As part of a series of studies describing the oculomotor behavior of skilled readers, a study investigated whether a word refixation curve exists. Subjects, 66 college students fixating over 40,000 times, read lines of text from a computer screen and were instructed to read for meaning without regard to errors. Results of eye movement control indicated that the frequency of immediately refixating a word following an initial eye fixation on it varies with the location of the fixation and that the data fit a parabolic function. The study concluded that the observed curvilinear relation results naturally from models that postulate a linear decrease in acuity with retinal eccentricity. Effects of word length and cultural frequency on the frequency of refixating were also examined. (Six figures and three tables of data are included; 27 references are attached.) (RS)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED307592

CENTER FOR THE STUDY OF READING

Technical Report No. 469

EYE MOVEMENT CONTROL DURING READING: II. FREQUENCY OF REFIXATING A WORD

G. W. McConkie, P. W. Kerr, M. D. Reddix, D. Zola
University of Illinois at Urbana-Champaign

A. Jacobs
Laboratoire de Psychologie Experimentale
Universite Rene Descartes

May 1989

University of Illinois at Urbana-Champaign
51 Gerty Drive
Champaign, Illinois 61820

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

D Anderson

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

CS005674

The work upon which this publication was based was supported in part by the Office of Educational Research and Improvement under Cooperative Agreement No. OEG 0087-C1001, with the Reading Research and Education Center and in part by Grant #HD18116 from the National Institute of Child Health and Human Development. The publication does not necessarily reflect the views of the agencies supporting the research.

BEST COPY AVAILABLE

EDITORIAL ADVISORY BOARD
1988-89

Beck, Diana

Commeyras, Michelle

Foertsch, Daniel

Hartman, Doug

Jacobson, Michael

Jehng, Jihn-Chang

Jimenez, Robert

Kerr, Bonnie

Kerr, Paul

Meyer, Jennifer

Moran, Juan

Ohtsuka, Keisuke

Roe, Mary

Schommer, Marlene

Scott, Judy

Stallman, Anne

Wilkinson, Ian

Wolff, Phillip

MANAGING EDITOR
Mary A. Foertsch

MANUSCRIPT PRODUCTION ASSISTANTS
Delores Plowman
Nancy Diedrich

Abstract

An analysis of over 40,000 eye fixations made by college students during reading indicates that the frequency of immediately refixating a word following an initial eye fixation on it varies with the location of that fixation. The refixation frequency is lowest near the center of the word and positively accelerating with distance from the center. The data are well fit by a parabolic function. Assuming that refixation frequency is related to the frequency of successful word identification, the observed curvilinear relation results naturally from models that postulate a linear decrease in acuity with retinal eccentricity. A single letter difference in fixation location in a word can make a sizeable difference in the likelihood of refixating that word. The effects of word length and cultural frequency on the frequency of refixating are also examined.

EYE MOVEMENT CONTROL DURING READING: II. FREQUENCY OF REFIXATING A WORD

During reading, people fixate more frequently near the centers of words than near the beginnings and ends (Rayner, 1979). O'Regan (1981) proposed a Convenient Viewing Position hypothesis, stating that readers learn to send their eyes to the centers of words because, across the words in a language, that location is optimal for word identification. Due to the rapid drop in visual acuity with distance from the center of the fovea, together with the fact that letters bounded by spaces are more perceptible (Bouma, 1978; Jacobs, 1987), the letters of a word are maximally identifiable when the eyes are near the word's center. O'Regan further suggested that the further the eyes lie from the Convenient Viewing Position, the greater the probability that a second eye fixation on the word will be required for identification. This prediction was confirmed in a word identification study (O'Regan, 1984), which found that the frequency of making a second fixation on a word is minimized when the initial fixation is near the center of that word. Furthermore, the frequency of refixating increases as the distance of the first fixation from the center of the word increases. We refer to this relationship as the *Word Refixation Frequency Curve*, or simply as the word refixation curve. The existence of a word refixation curve in word identification tasks has been replicated (O'Regan, Levy-Schoen, Pynte, & Brugaillere, 1984) and there is some evidence that it may also be present in the eye fixation pattern made during reading (Blanchard & McConkie, cited by O'Regan & Levy-Schoen, 1987).

The study reported below is the second in a series of studies describing the oculomotor behavior of skilled readers (see McConkie, Kerr, Reddix & Zola, 1988). Its purposes were: (a) to establish whether or not readers' eye movements show a word refixation curve; (b) to identify the characteristics of the curve, if it exists; (c) to determine whether or not the curve reflects word-level processes operating during reading or is a statistical artifact; and (d) to examine how two variables known to influence word perception, that is, word length and word frequency (Gough, 1984), affect the shape of the curve, thus clarifying their role in the ongoing reading process.

It would be reasonable to expect that, while a word refixation curve occurs in tasks involving identification of isolated words, it may not be found in the reading of connected discourse. First, the contextual constraint in normal text may permit words to be typically identified on a single eye fixation, with refixations being required only on very long words. Second, peripherally-acquired information about subsequent words may reduce or eliminate the need for refixations, even though they may be required in identifying isolated words. Third, there may be insufficient time within individual eye fixations in normal reading to determine whether or not the fixated word has been successfully identified and to accordingly adjust the next saccade.

If a characteristic word refixation curve is present in the eye movement behavior of skilled readers, it could still simply be a statistical artifact. The frequency distribution of lengths of saccades made during reading is bimodal, with one mode for forward saccades, and a second for regressive saccades. The frequency of short saccades of less than 2-3 character positions in length is quite low. This being the case, a model that assumes that during each eye fixation the following saccade is selected randomly from this distribution would produce an artifactual word refixation curve. When the eyes are centered near the middle of a word, more of the letters of that word lie within the central part of the saccade length distribution, where their frequency of being selected for the next saccade is low. In order to test the claim that a word refixation curve during reading results from word identification processes, it is necessary to show that the obtained curve differs from that predicted by a random selection model of the type just described.

Assuming that a non-artifactual word refixation curve is found in normal reading data, and that it results from characteristics of vision and word identification in the manner postulated by O'Regan, then the curve should show effects of word length and word frequency. Longer words extend further into the periphery and thus contain letters of lower perceptibility than do shorter words. This should result in more refixations of longer words when the eyes are the same distance from the center of the word. More common words are more easily identified, that is, they require less complete visual information for identification (Howes & Solomon, 1951; Postman & Adis-Castro, 1957), and this should reduce the need for refixations. These predictions were tested in the study.

Methods

Data were taken from a series of studies conducted by McConkie, Reddix & Zola (1985). Sixty-six college students read the first two chapters (994 lines of up to 73 characters per line) of a contemporary novel from a computer screen. Four characters subtended one degree of visual angle. Eye movement data were collected while reading 300 of the lines, using an SRI Dual Purkinje Image Eyetracker (Cornsweet & Crane, 1973) sampling eye position every ms. For purposes unrelated to the current study, in 20-25% of the lines of text a single word was replaced by either a non-word letter string or an erroneous word. Subjects were instructed to read the passage for meaning, without regard to the errors. They were tested for comprehension at 15 locations, and were never asked about the errors.

For each subject, either 50 or 100 of the lines on which data were collected contained no errors. Data from these control condition lines, consisting of a total of 43,668 eye fixations, were used in the analyses described below.

Results

Description of the word refixation curve. Data were selected in the following way: Only the initial fixation on a word was included in the analysis, and then only if it was preceded by a forward saccade. Cases were excluded if the reader's eyes had ever fixated to the right of the word on the current line, or if the fixated word was immediately preceded or followed by punctuation. Finally, first and last fixations on a line and other cases where the eyetracker's signal was disturbed (e.g., a blink) were not used in this analysis.

Figure 1 presents the proportion of initial fixations at different letter positions on words of different lengths that were immediately followed by a refixation on the word. The proportions of refixations for each word length manifest a word refixation curve that is well-specified by the equation:

$$Y = A + B (X - C)^2$$

In this equation, the parameter *A* indicates the height or vertical offset of the curve at its lowest point, *B* indicates the slope of the curve, and *C* indicates the letter position in the word where the curve reaches its lowest point, that is, the horizontal offset of the curve. Table 1 contains estimates of these three parameters for words of length 4 through 8, as a result of fitting the curve using the PAR program in the BMDP statistical package (Dixon, 1981). Adjacent parameter values that do not differ by at least two standard deviations are joined by a vertical bar.

[Insert Figure 1 & Table 1 about here.]

Influence of word length. Word length has its greatest effect on the parameter *C*, indicating that the minimum point in the curve shifts rightward as word length increases. However, the difference between *C* and the middle of the word is small and nearly constant across word lengths (see Table 1). The center of the second letter of a word is designated as position 2.5 in the data. Therefore, in Table 1 (Column 9) the center of a 4-letter word is position 3.0, which is the midpoint between the second and third letters. For 4- to 8-letter words, the optimal fixation location, defined as the location where an immediate refixation of that word is least likely to occur during reading, is 1/4 to 1/2 character position to the right of the center of the word (Table 1, Column 10). While small, this difference is statistically significant: for each word length, the value of C minus center is more than two standard errors above zero.

There is a tendency for the minimum point of the word refixation curve (i.e., parameter *A*) to increase with word length. The exception, 4-letter words, will be discussed later. This increase indicates that there is a somewhat greater frequency of refixating a longer word than a shorter word, following a fixation at a corresponding distance from the center of the word. Parameter *B*, the slope parameter, is very similar for words of different lengths. Except for 4-letter words, there is no influence of word length on the slope.

To illustrate the similarity of the curves for words of different lengths, the data points for words of each length were adjusted by parameters *A* and *C* (i.e., the value for parameter *C* for each word length was subtracted from the X values of the data points, and the value for parameter *A* was subtracted from the Y values). The resulting data points are plotted in Figure 2, together with a curve (dotted line) having a *B* parameter value of .03. As can be seen, word length appears to have no influence on the slope of the refixation curve, and the entire set of data is fit very well by the proposed function.

[Insert Figure 2 about here.]

Testing whether the curve is an artifact. In order to determine whether the obtained word refixation curve should be considered artifactual, predicted word refixation curves for words of different lengths were derived from the model of random eye movement control described above. To do this, an overall frequency distribution of saccade lengths was obtained from the data set (see Figure 3). Then a predicted word refixation curve for words of each length, from 2 to 10 letters, was obtained in the following manner. Assuming words of a given length, the overall distribution was first centered on the space at the left of the word. With the distribution at that location, the proportion of the overall distribution that lay within the region occupied by the word (including the space lying immediately to the left of the word) was calculated (see Figure 3). According to the model, this indicates the proportion of times that an initial fixation on the space at the left of the word will be followed by a refixation on the word if saccades are randomly selected from the overall distribution. This proportion was assigned to that space as the *predicted refixation frequency* for that character position. This process was then repeated for each letter position in the word, giving a predicted refixation frequency for each position. Finally, the predicted refixation frequency values for each character position in words of a given length were plotted, together with the obtained proportion of refixations following fixations at these locations. Figure 4 shows these curves for words of four different lengths.

[Insert Figures 3 & 4 about here.]

As Figure 4 indicates, the predicted word refixation curves for shorter words show some characteristics of the obtained curves: they are somewhat U-shaped though the location of the minimum varies with word length. However, the fit is relatively poor. In a second attempt at testing the random control model, separate frequency distributions were prepared for fixations on words of different lengths. Predicted word refixation curves were then created for words of each length using only data for saccades originating from words of that length. The results were very similar to those presented in Figure 4. Hence, the random model, by itself, cannot account for the data patterns obtained here.

The obtained word refixation curves differ from the predicted curves in two ways. First, the obtained curves are much steeper, with the frequency of refixation being considerably higher than the predicted values at the beginning of the words and lower at the center. Second, in the obtained data, the minimum point is near the center of the word for all word lengths while this is not true for the predicted curves. Both of these differences indicate that the frequency of refixation is being heavily influenced by the location of the eyes in a word, rather than simply being the result of the random selection of saccade lengths. Of course, these data cannot refute the possibility that the lengths of some saccades are selected randomly without direct word-based influences but we know of no way to test such a possibility.

Influence of word frequency. In order to examine the influence of word frequency on the word refixation curve, each initial fixation on a word was associated with the log of the sum of one plus the cultural frequency of the fixated word (Kucera & Francis, 1967). Fixations were then partitioned into four groups, depending on the log frequency of the fixated word: 0 to 0.99, 1.0 to 1.99, 2.0 to 2.99, and 3.0 and above. Figure 5 presents the word refixation curves for words in each of these frequency bands, combined across word lengths, and Table 2 presents the parameters obtained by fitting the proposed function to each of these sets of data. In fitting the function to the data, only data points based on at least 100 fixations were included. Data points based on fewer than 100 fixations are presented in Figure 5, but with their symbols circled to distinguish them from the more stable data points. In Table 2, adjacent parameter values that do not differ by at least two standard deviations are joined by a vertical bar.

[Insert Table 2 & Figure 5 about here.]

Word refixation curves for log frequencies of 0 through 2.99 are quite similar in shape, though vertically offset from one another. As frequency decreases, the vertical offset (parameter A) increases, indicating a greater frequency of refixation of lower frequency words. Little difference is observed in the slope of the curve (parameter B), or in the horizontal offset of the curve (parameter C).

The data from the highest frequency words consists entirely of fixations on 4-letter words, most being function words; in the passage used, no words of length 5 through 8 fall into this frequency band. The word refixation curve for these data is somewhat different from the curves for lower frequency words, being less steep, and showing a minimum further to the right of the center of the word.

Thus, the proposed function fits the data from all frequency bands quite well. With the exception of high frequency 4-letter words, the primary effect of word frequency is to influence the height of the word refixation curve.

In the analyses described above, both word length and word frequency have been shown to produce an effect on A, the vertical offset parameter. However, there are two reasons for why this effect might be due to only one of these two variables. First, the two variables are correlated. In the current data set, the correlation between the length and log frequency of fixated words is -.46. Second, the combining of data across one variable (i.e., word length) in order to examine the effect of the other (i.e., word frequency) invites confounding of these variables. In particular, fixations at locations more distant from the center of the word are found only in longer words, hence changing the amount of influence of words of different lengths at the different letter positions in the word.

In order to determine whether the effects observed result from only one of these two variables, it is necessary to partition the data by word length and word frequency simultaneously, and examine the word refixation curves for the resulting cells. However, partitioning the data in this manner produced data sets for many cells that were too small to yield stable refixation frequency estimates for the different letter positions. Therefore, a formal analysis cannot be carried out to answer the question. However, Figure 6 presents word refixation curves for some cells where sufficient data exist. Holding word frequency constant at 1.0-1.99, Figure 6a presents word refixation frequency curves for words of the lowest three frequency bands. Data points are not included where fewer than 50 fixations were made on a given letter position in the word. An examination of these figures indicates that when word frequency is held constant, there is little vertical separation among the curves for words of different length. Varying word frequency with word length held constant, however, does produce a difference in vertical separation especially when the eyes are not at the center of the word. Thus, it appears likely that the variation in vertical offset, the A parameter, is primarily due to differences in word frequency rather than to differences in word length, though this requires closer examination with a larger data set.

[Insert Figure 6 about here.]

Discussion

Previous studies using word recognition tasks have found that the frequency of refixating a word is a function of the location of the initial fixation location on the word, with fewest refixations following a fixation at the center of the word. The results of the present study have demonstrated that this same phenomenon occurs in normal reading. Furthermore, the relationship between the initial fixation on

a word and the frequency of refixating it takes the shape of a parabola. This Word Refixation Frequency Curve is well fit by the equation:

$$Y = A + B * (X - C)^2.$$

In this equation, *X* is the letter position on which a reader's eyes are centered during an initial fixation on the word, and *Y* is the proportion of the immediately following fixations in which the eyes are centered on the same word. The parameter *A*, the height or *vertical offset parameter*, is the height of the curve at its lowest point. This is the frequency that a fixation made at the optimal location in the word will be followed by a refixation. *B*, the *slope parameter*, indicates the steepness of the rise in refixation frequency as the location of the initial fixation deviates from the optimal location. This indicates the penalty paid for not fixating the center of the word. The rise is not linear, but shows positive acceleration with distance from the minimum. *C*, the *horizontal offset parameter*, is the letter position in the word where an initial fixation is least likely to be followed by a refixation.

The value of parameter *C* is consistently within 1/4 to 1/2 character position to the right of center for words of the lengths studied here. Thus, if eye position, *X*, is measured from the center of the word rather than from its beginning, it may be possible to make this parameter a constant. At present, our best estimate for the value of *C* minus center is .41. This would reduce the number of free parameters in the equation from three to two.

The results from this study are consistent with O'Regan's (1984) Convenient Viewing Position Hypothesis, which indicates that there is an optimal location within each word from which identification can occur most readily, that the average optimal location is near the center of the word, and that a penalty is paid on fixations deviating from that location. While O'Regan reported this penalty in terms of average gaze duration, the present results indicate the penalty in terms of an increased frequency of refixating. On the average, making a second fixation on a word effectively doubles the gaze duration on it. Thus, a difference of 30 msec in gaze duration (or 30 msec penalty in O'Regan's terms) could be the result of raising the frequency of refixating from 5% to 20%, rather than being due to any change in individual fixation durations.

For people involved in reading research, the most important finding from the current study is that the word refixation curve is surprisingly steep. While the frequency of refixating a 7-letter word when the eyes are at their optimal location is .09, this rises to .20 when the eyes are only two letters away from that location and .34 when three letters away. Thus, one letter position difference in where the eyes land in a word can make a substantial difference in the frequency of refixation. We assume that this results from the rapid drop in visual acuity as a function of distance from the center of the fovea.

With O'Regan, we assume that a majority of refixations result from failure to identify a word by the time the following saccade is requested. This is consistent with the observed relationship between fixation location and refixation frequency. However, some refixations are probably due to difficulties at higher processing levels, thus keeping the eyes from advancing to the next word until the reader is ready for new visual input. Other refixations probably result from error in the perceptuo-oculomotor system, which sometimes sends the eyes short of the targeted word, thus producing an inadvertent refixation (McConkie, Kerr, Reddix & Zola, 1988). At the same time, word identification failure does not necessarily produce an immediate refixation. For example, replacing a word with a pseudo-word or with a letter string that violates English spelling patterns increases the frequency of refixations on that word, but most of the time readers still fixate the word only once (McConkie, Reddix & Zola, in preparation). Thus, the frequency of making an immediate refixation. For example, replacing a word with a pseudo-word or with a letter string that violates English spelling patterns increases the frequency of refixations on that word, but most of the time readers still fixate the word only once (McConkie, Reddix & Zola, in preparation). Thus, the frequency of making an immediate refixation relates to, but does not directly indicate, the frequency of initial word identification failure during reading.

The assumption that refixation frequency reflects the frequency of initial word identification failure leads to a further question. Since visual acuity decreases as a linear function of the distance from the center of vision (Olzak & Thomas, 1986), why does the refixation frequency increase as an accelerating function of this distance? An answer to this question is given by a simple summed letter information model of

word identification during fixations in reading. This model makes three assumptions: (a) the amount of visual information obtained from a letter decreases as a linear function of its distance from the center of vision, (b) the total amount of visual information available from a word is the sum of the information available from all its letters, and (c) the frequency of identifying a word during the initial fixation on it is a linear function of the amount of visual information available from it. The third assumption is qualified by assuming a high threshold above which a word is always identified, and a low threshold below which it is not identified. For present purposes we will ignore a fourth assumption, that the distance between adjacent letters influences their perceptibility, since it makes no real difference in the following derivation.

Table 3 presents an example of an application of this simple model. It shows the amount of information available from individual letters in a seven-letter word, and from the total word, based on the assumptions that the amount of information from the directly fixated letter has the value of 1, and that there is a drop of .1 (an arbitrarily chosen value) in information for each letter position unit of distance from that location. Jacobs (1979) estimates the rate of increase in the minimal angle of resolution, as a function of retinal eccentricity, to be .61 min of arc for each degree increase in eccentricity, for laterally masked letters. Given these assumed parameters, word identification failure is lowest with the eyes at the center of the word (e.g., total visual information is greatest with the eyes at that point), and the function relating fixation location to total visual information from a word is a parabola. Thus, a system having a linear decrease in acuity gives rise naturally to a parabolic decrease in the frequency of word identification failure with distance of the eyes from the center of the word. This is hypothesized to be the primary basis for the Word Refixation Frequency Curve.

[Insert Table 3 about here.]

Two variables were identified that affect the Word Refixation Frequency Curve. One variable, word length, influences the height and offset parameters, *A* and *C*, but has no effect on the slope parameter, B. The effect on the offset parameter has already been explained as resulting from the fact that the minimum refixation location is just to the right of the center of the word. The lack of an effect on the slope parameter, *B*, and the symmetry of the curve are unexpected. Other studies of word perception have demonstrated that words are identified better in the right visual field than in the left (e.g., Bouma, 1973). There is ample evidence that the perceptual span during the reading of English text extends further to the right than to the left (McConkie & Rayner, 1976; Rayner, Well, & Pollatsek, 1980; Underwood & McConkie, 1985). Letters lying more than about 4 to the left of the center of vision, and about 8 to the right, do not appear to be utilized during a fixation. If letters in these regions were assumed to provide no visual information for word identification in the model described above, an asymmetry in the word refixation curve is predicted that does not appear in the data. We have no explanation for the lack of the expected asymmetry, but this finding indicates the need for further study of the reasons for asymmetry in the perceptual span during reading.

The height parameter, *A*, tends to increase with word length. The frequency of refixation tends to be greater in longer words than in shorter words, at any fixation location. This is just opposite the prediction of a simple letter information model such as that described above. Such a model would typically add more information for each letter of the word and thus give more total visual information for longer words. However, since more long words exist in the lexicon, and since they are visually more complex than shorter words, it is reasonable to assume that more visual information is required in order to discriminate among them (Morton, 1969). Apparently this trade-off between amount of information gained from longer words, and amount required for their discrimination, leads to a total net loss in discriminability as length increases.

The second variable influencing the Refixation Frequency Curve is word frequency. Differences in word frequency have their primary effect on the height parameter, *A*. If the words in the highest frequency band, which are all 4-letter words occurring over 1,000 times per million, are excluded, then there is no statistically significant difference between either slope (B) or horizontal offset (if measured from the center of the word, i. e., *C* minus word center) parameters. Thus, higher frequency words are somewhat less likely to be refixated at all letter positions.

The lack of an effect of word frequency on the B (slope) parameter is somewhat surprising. Recent theories of reading have emphasized the role of word frequency in word identification, especially under incomplete visual stimulus conditions (see Gough, 1984 for a recent review). Thus, we might expect that as the initial fixation location moves away from the center of the word, reducing the amount of clear visual information provided by the word, the influence of word frequency would become greater. This should show up in the data as a lower slope parameter for higher frequency words. The data show a trend in this direction, but it is small and not statistically significant for any but the highest frequency category.

The Word Refixation Curve for short, high-frequency words was different than that of other word classes, being less steep and with a minimum slightly further to the right. It may also be that, for these words, the Refixation Curve is not symmetric, as it is for words of other frequencies. Finally, the curve does not show a lower vertical offset than do words of other frequencies, as might be expected. However, these data raise a more fundamental question. Why is it that high frequency, 4-letter words are ever refixated, and more particularly, why are they sometimes refixated following an initial eye fixation at their center? It seems unlikely that this is due to a failure to gain enough visual information from the word to identify it. If this is true, then we must conclude that the minimum refixation frequency in this condition indicates a frequency with which skilled readers refixate a word for reasons other than identification failure. The value of parameter *A* for these words is .08. This is quite similar to the values for words of all but the lowest frequency band, ranging from .06 to .08. Thus, we suggest that about 8% of the time, an initial fixation on a word is followed by a refixation produced by causes other than word recognition failure. Of course, this may vary with reading strategy and reading ability level.

What is it that might produce such refixations? It is not likely that they are instances in which the eyes were being sent to the following word but undershot it. With the eyes near the center of a short word, the center of the next word would usually be of a distance less than that of an average saccade. In this situation, the eyes tend to overshoot their target, rather than to undershoot it (Kapoula & Robinson, 1986). Thus, the cause may lie in the cognitive domain rather than in the perceptual or oculomotor. If so, the value of the *A* (vertical offset) parameter for short, high frequency words may be of psychological interest, indicating the frequency with which refixations occur on some basis other than unavailability of visual information, for a given subject or a given text or reading condition. Further research is needed to investigate the actual basis for such refixations. Possibilities include conceptual difficulties in the message, processing lagging behind the eye fixation pattern, or a previous overshooting of the target word.

In summary, data on the frequency of refixating a word, given an initial eye fixation at different letter positions in it, are fit well by a three parameter model. The *C* parameter, indicating horizontal offset, may be fixed at about .4 of a character position to the right of the center of a word, this apparently being the optimal position in a word for its identification. The *A* parameter, vertical offset, indicates the frequency of refixation when the eyes are at their optimal position in the word. When this parameter is obtained for fixations on short, high frequency words, it probably indicates the frequency of refixations that occur for reasons other than insufficient visual information. The *B* parameter, slope, probably indicates the rate of dropoff of visual information necessary for reading as a function of retinal eccentricity. Thus, it may vary with visibility of the text. It may also vary with reading strategy, if readers can change their degree of carefulness in reading, thereby changing the amount of visual information required for word identification.

Word length and word frequency affect only parameter *A* (if letter position is measured from the word's center) in all but short, high frequency words. However, these two variables are correlated, and the data set in the current study was not large enough to test whether the effect resulted from only one of these variables. Available evidence suggests that the effect is primarily due to variation in word frequency. Short, high frequency words have a refixation curve that is flatter than that of words in other frequency bands, indicating better identification at peripheral locations. This is consistent with reports of skipping the word "the" during reading (O'Regan, 1979, 1980; Rayner, 1977).

There is one final implication of the findings from the current study that should be noted. When skilled readers move their eyes to a new word, they appear to be sending them to a functional target location

at or near the center of that word (McConkie, Kerr, Reddix & Zola, 1988), the location at which refixations are the least frequent. However, there are two sources of error that cause the eyes to deviate from that location: (a) nonsystematic error that results in a normally-distributed collection of landing sites, and (b) a systematic saccadic range error by which the eyes overshoot or undershoot the functional target depending on the location of the preceding fixation (Kapoula & Robinson, 1986). The refixation curve observed in the present study suggests that these errors in eye placement have a substantial negative effect on reading time. The effect of both sources of error is to reduce the frequency with which the eyes land at the center of the targeted word. This, in turn, increases the frequency with which refixations are required, thus increasing the reading time. The size of this increase depends not only on the increased frequency of refixations, but also on the durations of these additional eye fixations. Factors influencing the durations of eye fixations will be the topic of another report in this series.

Of course, any visual or oculomotor abnormality that either reduces visual acuity or reduces the frequency of placing the eyes at the centers of targeted words should produce an increase in the frequency of refixating, further slowing the reading rate. For children learning to read, such abnormalities would reduce the visual information typically available from a word during eye fixations. What effects this might have on the learning process is an important matter for future research.

References

Bouma, H. (1973). Visual interference in the parafoveal recognition of initial and final letters of words. *Vision Research, 13*, 767-782.

Bouma, H. (1978). Visual search and reading: Eye movements and functional visual field: A tutorial review. In J. Requin (Ed.), *Attention and Performance VII*. New York: Erlbaum.

Cornsweet, T. N., & Crane, H. D. (1973). Accurate two-dimensional eye tracker using first and fourth Purkinje images. *Journal of the Optical Society of America, 63*, 6-13.

Dixon, W. J. (1981). *BMDP Statistical Software*. Berkeley: University of California Press.

Gough, P. B. (1984). Word recognition. In P. D. Pearson (Ed.), *Handbook of Reading Research*. New York: Longman.

Howes, D. H., & Solomon, R. L. (1951). Visual duration threshold as a function of word-probability. *Journal of Experimental Psychology*, 41, 401-410.

Jacobs, R. J. (1979). Visual resolution and contour interaction in the fovea and periphery. *Vision Research, 19*, 1187-1195.

Jacobs, A. M. (1987). On localization and saccade programming. *Vision Research, 27*(11), 19593-1966.

Kapoula, Z., & Robinson, D. A. (1986). Saccadic undershoot is not inevitable: Saccades can be accurate. *Vision Research, 26*(5), 735-743.

Kucera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.

McConkie, G. W., & Rayner, K. (1976). Asymmetry of the perceptual span in reading. *Bulletin of the Psychonomic Society, 8*(5), 365-368.

McConkie, G. W., Kerr, P. W., Reddix, M. D., & Zola, D. (1988). Eye movement control during reading: I. The location of initial fixations on words. *Vision Research, 28*(10), 1107-1118.

McConkie, G. W., Reddix, M., & Zola, D. (1985). *Chronometric analysis of language processing during eye fixation in reading*. Paper presented at the Annual Meeting of the Psychonomic Society, Boston, MA.

McConkie, G. W., Reddix, M. D., & Zola, D. (in preparation). *Tracing the time course of language processing during eye fixations in reading*.

Morton, J. (1969). Interaction of information in word recognition. *Psychological Review, 76*, 165-178.

Olzak, L. A., & Thomas, J. P. (1986). Seeing spatial patterns. In K. R. Boff, L. Kaufman, & J. P. Thomas Eds.), *Handbook of perception and human performance: Vol. 1: Sensory processes and perception* (pp. 7.1-7.56). New York: John Wiley.

O'Regan, K. (1979). Saccade size control in reading: Evidence for the linguistic control hypothesis. *Perception and Psychophysics, 25*, 501-509.

O'Regan, K. (1980). The control of saccade size and fixation duration in reading: The limits of linguistic control. *Perception and Psychophysics, 28*, 112-117.

O'Regan, K. (1981). The "convenient viewing position" hypothesis. In D. F. Fisher, R. W. Monty, & J. W. Senders (Eds.), *Eye movements: Cognition and visual perception* (pp. 289-298). Hillsdale, NJ: Erlbaum.

O'Regan, J. K. (1984). How the eye scans isolated words. In A. G. Gale & F. Johnson (Eds.), *Theoretical and applied aspects of eye movement research.* Amsterdam: North-Holland.

O'Regan, K., & Levy-Schoen, A. (1987). Eye-movement strategy and tactics in word recognition and reading. In M. Colheart (Ed.), *Attention and Performance XII* (pp. 363-384). Hillsdale, NJ: Erlbaum.

O'Regan, J. K., Levy-Schoen, A., Pynte, J., & Brugaillere, B. (1984). Convenient fixation location within isolated words of different length and structure. *Journal of Experimental Psychology: Human Perception and Performance, 10*, 250-257.

Postman, L., & Adis-Castro, G. (1957). Psychophysical methods in the study of word recognition. *Science, 125*, 193-194.

Rayner, K. (1977). Visual attention in reading: Eye movements reflect cognitive processing. *Memory & Cognition, 4*, 443-448.

Rayner, K. (1979). Eye guidance in reading: Fixation locations with words. *Perception, 8*, 21-30.

Rayner, K., Well, A. D., & Pollatsek, A. (1980). Asymmetry of the effective visual field in reading. *Perception & Psychophysics, 27*, 537-544.

Underwood, N. R., & McConkie, G. W. (1985). Perceptual span for letter distinctions during reading. *Reading Research Quarterly, 20*, 153-162.

Authors Note

Copies of this paper can be obtained from George W. McConkie, Center for the Study of Reading, 51 Gerty Drive, Champaign, IL 61820.

Table 1

Estimated Parameters A, B, and C for Words of Length 4 Through 8[1]

Word Length	Parameters A[3]	Parameters B[3]	Parameters C[3]	Average Residual	Sum of Squared Residuals	Total N	Number of Refixations	Center of Word[2]	C Minus Word Center[2,3]
4	.074 (.002)	.025 (.001)	3.24 (.037)	.0001	.021	4291	564	3.0	.24
5	.055 (.007)	.028 (.002)	3.93 (.088)	.0002	.296	3797	564	3.5	.43
6	.082 (.010)	.032 (.002)	4.52 (.103)	.0018	.735	3325	750	4.0	.52
7	.088 (.016)	.029 (.003)	4.99 (.160)	.0020	1.768	2384	577	4.5	.49
8	.134 (.008)	.028 (.001)	5.37 (.038)	.0015	.262	1205	375	5.0	.37

Note: Standard deviations of parameters are in parentheses.

Function: $Y = A + B * (X - C)^2$

Y = Proportion of initial fixations that were immediately followed by a refixation of the word
X = Location of initial fixations in word (character position)
A = Vertical Offset
B = Slope
C = Horizontal Offset

[1]Fixations on the space to the left of a word were associated with that word. Thus, data on a five-letter word included fixations on six letter positions.

[2]Measured in character position units.

[3]Adjacent values that do not differ by at least 2 standard errors are joined by a vertical bar.

Table 2

Estimated Parameters A, B, and C for Four Word Frequency Groups

Frequency Group	Parameters A[1]	Parameters B[1]	C Minus Word Center[1]	Average Residual	Sum of Squared Residuals	Total N	Number of Refixations
0.0-0.9	.122 (.010)	.031 (.002)	.406 (.128)	.0004	.617	2780	739
1.0-1.9	.074 (.011)	.031 (.002)	.379 (.112)	-.0044	1.897	5409	1072
2.0-2.9	.057 (.011)	.027 (.003)	.416 (.154)	.0050	1.496	5469	839
3.0+	.038 (.011)	.013 (.004)	.887 (.422)	.0002	.262	1344	180

Note: Standard deviations of parameters are in parentheses.

[1]Adjacent values that do not differ by at least 2 standard errors are joined by a vertical bar.

Table 3

A Linear Decrease in Letter Information With Retinal Eccentricity Produces a Nonlinear Relation Between Fixation Location and Word Information

Letter Position Fixated in Word	Position of letter in a 7-letter Word: 1	2	3	4	5	6	7	Total Visual Information from Word
	Visual Information from Individual Letters							
0	.9	.8	.7	.6	.5	.4	.3	4.2
1	1	.9	.8	.7	.6	.5	.4	4.9
2	.9	1	.9	.8	.7	.6	.5	5.4
3	.8	.9	1	.9	.8	.7	.6	5.7
4	.7	.8	.9	1	.9	.8	.7	5.8
5	.6	.7	.8	.9	1	.9	.8	5.7
6	.5	.6	.7	.8	.9	1	.9	5.4
7	.4	.5	.6	.7	.8	.9	1	4.9

Note: Letter position 0 is the space to the left of the word.

Assumptions: Amount of visual information from directly fixated letter is equal to 1. This information drops by .1 for each letter position of eccentricity. Word information is the sum of letter information.

Figure Captions

Figure 1. Proportion of initial fixations at different locations in a word that were immediately followed by a refixation on the same word. Letter position 0 is the space to the left of the word. Proportions are based on sample sizes of 1205 to 4291 eye fixations. Refixations are least frequent when the initial fixation is near the center of the word.

Figure 2. Proportion of initial fixations at different locations on a word that were immediately followed by a refixation on the same word. The value for parameter *C* (horizontal offset) for each word length has been subtracted from the X values of the data points, and the value for parameter *A* (vertical offset) has been subtracted from the Y values. Word length (range: 4-8) has no influence on the slope of the refixation curve. The data are well fit by the proposed function, with B=.03, as represented by the dotted line.

Figure 3. A frequency distribution of the directions and lengths of saccades in the data set, regardless of word length. The shaded portion of the distribution represents the proportion of saccades that would be counted as refixations if the eyes were centered on the space to the left of a five-letter word, and the next saccade were selected randomly from the saccade distribution.

Figure 4. Observed refixation frequency curves plotted against predicted curves derived from a model of random eye movement control for words of different lengths. The random model provides a poor fit to the data.

Figure 5. Proportion of refixations following initial fixations at different locations in a word for words in different frequency bands. Letter position is measured from the center of the word. For example since 3.5 is the center of a 5-letter word, data for letter position 0 are based on fixations lying between 3.0 (the left edge of the third letter in the word) to 3.9. Since the center of a 4-letter word is 3.0, data for letter position 0 are based on fixations lying between 2.5 (the center of the second letter) and 3.4. Circled data points are based on fewer than 100 data values.

Figure 6. Refixation frequency curves with word frequency or word length held constant. Figure 6a shows refixation frequency curves from words of different lengths, but all having a cultural frequency of 10 to 999 per million (log frequency of 1.0 to 1.9). Figure 6b shows refixation curves for 6-letter words of different cultural frequency. It appears that variation in the vertical offset parameter, A, is primarily due to word frequency differences.

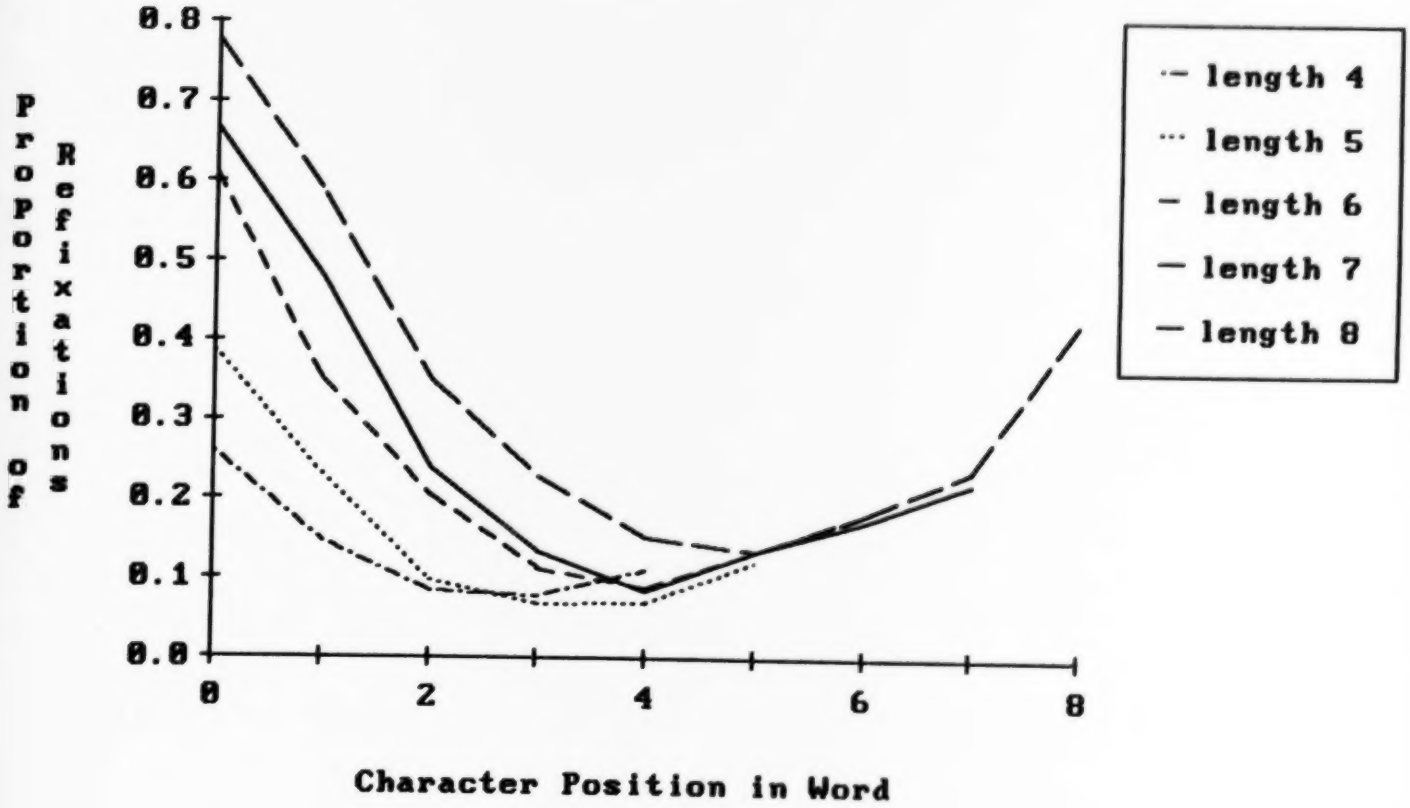

Proportion of Refixations
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0
2
4
6
8
Character Position in Word
length 4
length 5
length 6
length 7
length 8

Figure 1

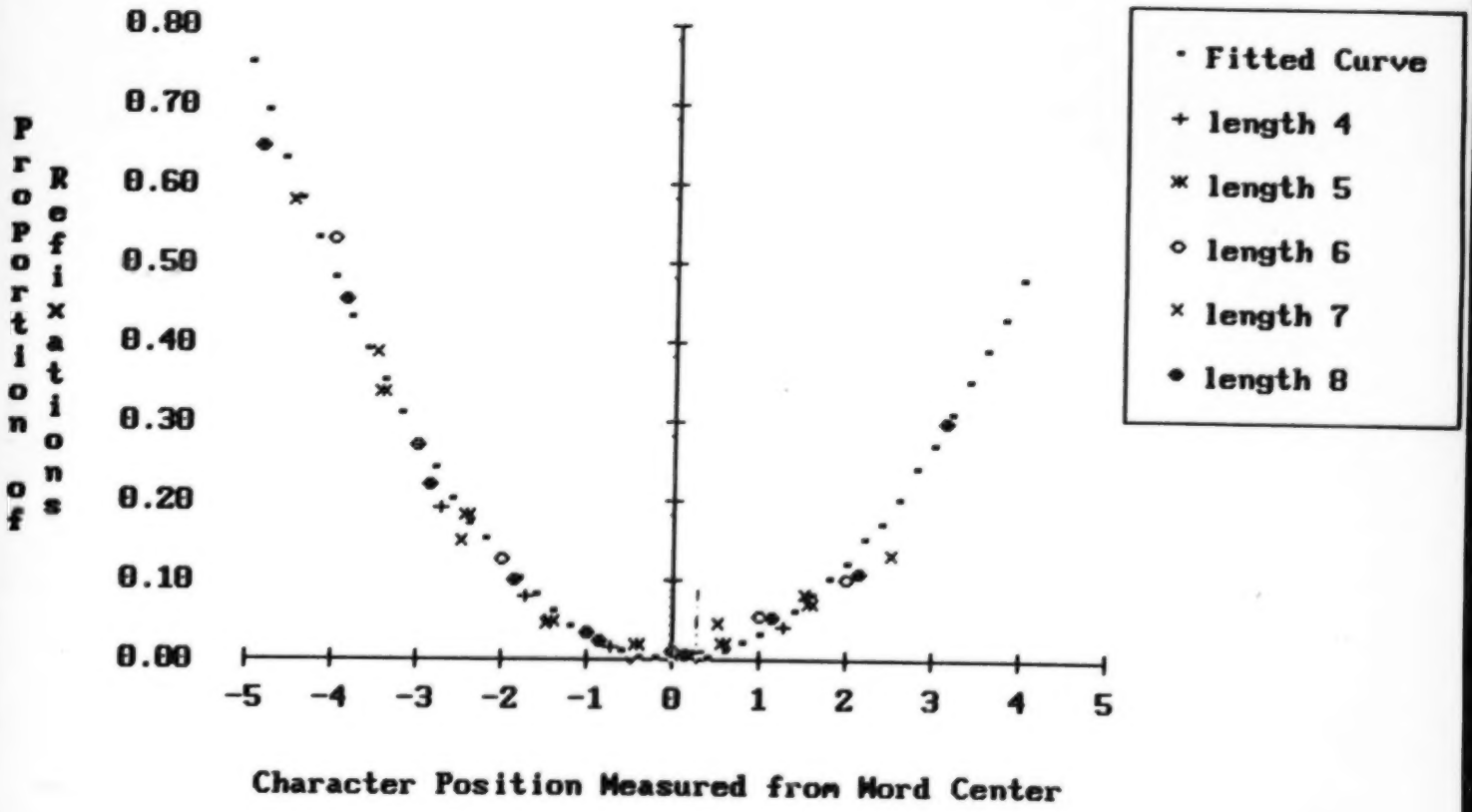

Proportion of Refixations
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
-5
-4
-3
-2
-1
0
1
2
3
4
5
Character Position Measured from Word Center
Fitted Curve
length 4
length 5
length 6
length 7
length 8

Figure 2

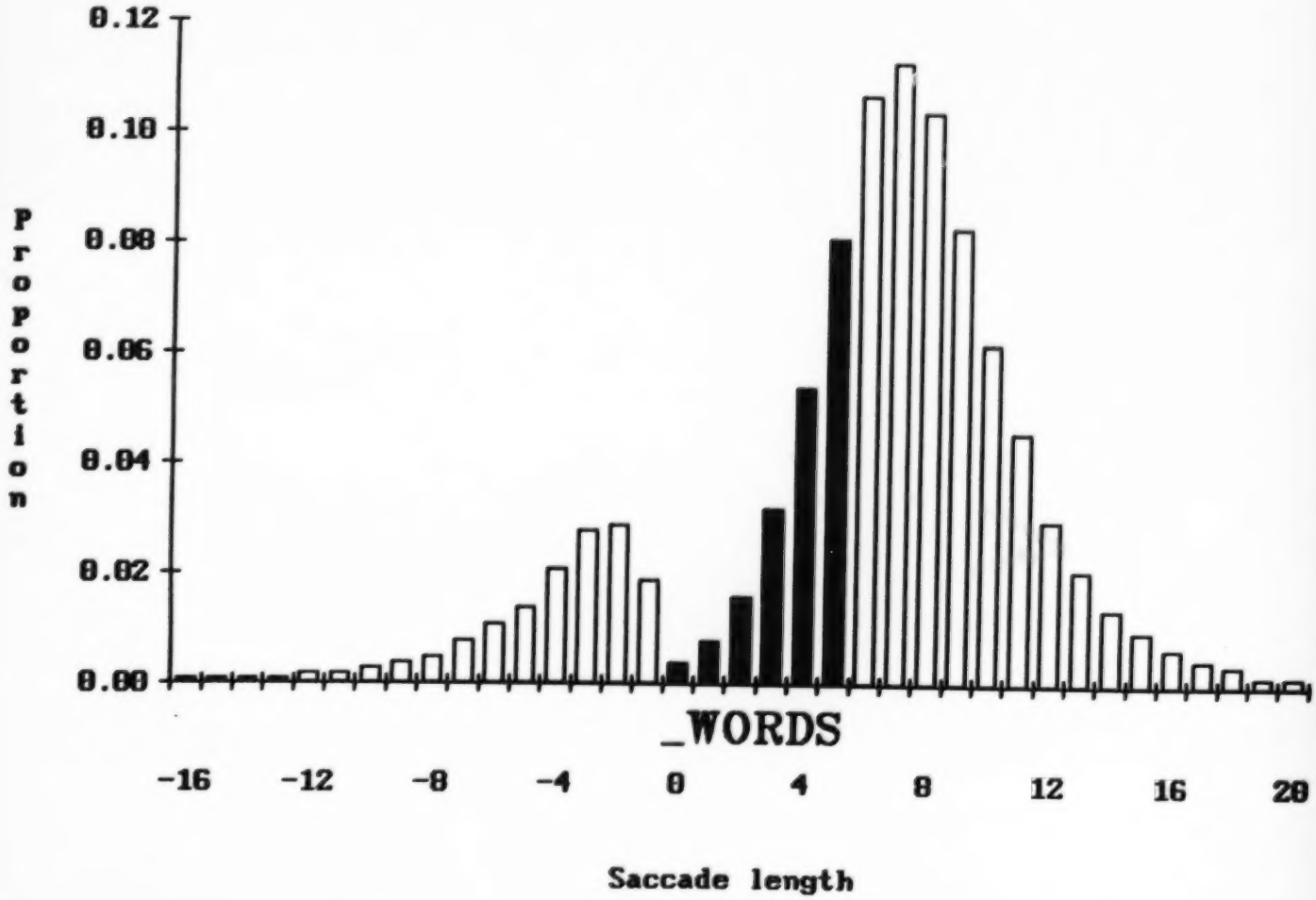

0.12
0.10
0.08
0.06
0.04
0.02
0.00
Proportion
-16
-12
-8
-4
0
4
8
12
16
20
_WORDS
Saccade length

Figure 3

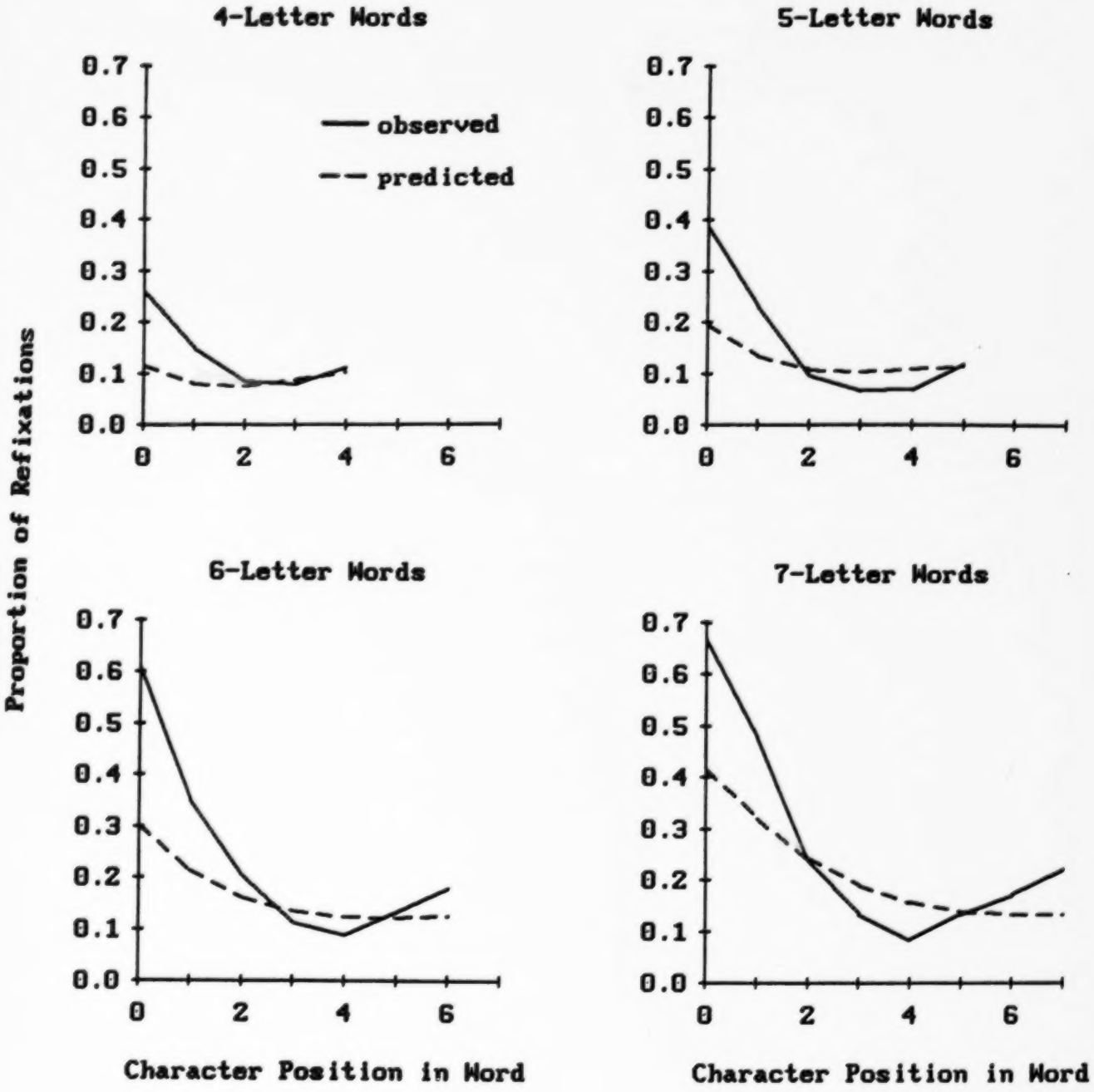

4-Letter Words
5-Letter Words
6-Letter Words
7-Letter Words
observed
predicted
Proportion of Refixations
Character Position in Word
Character Position in Word
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0
2
4
6

Figure 4

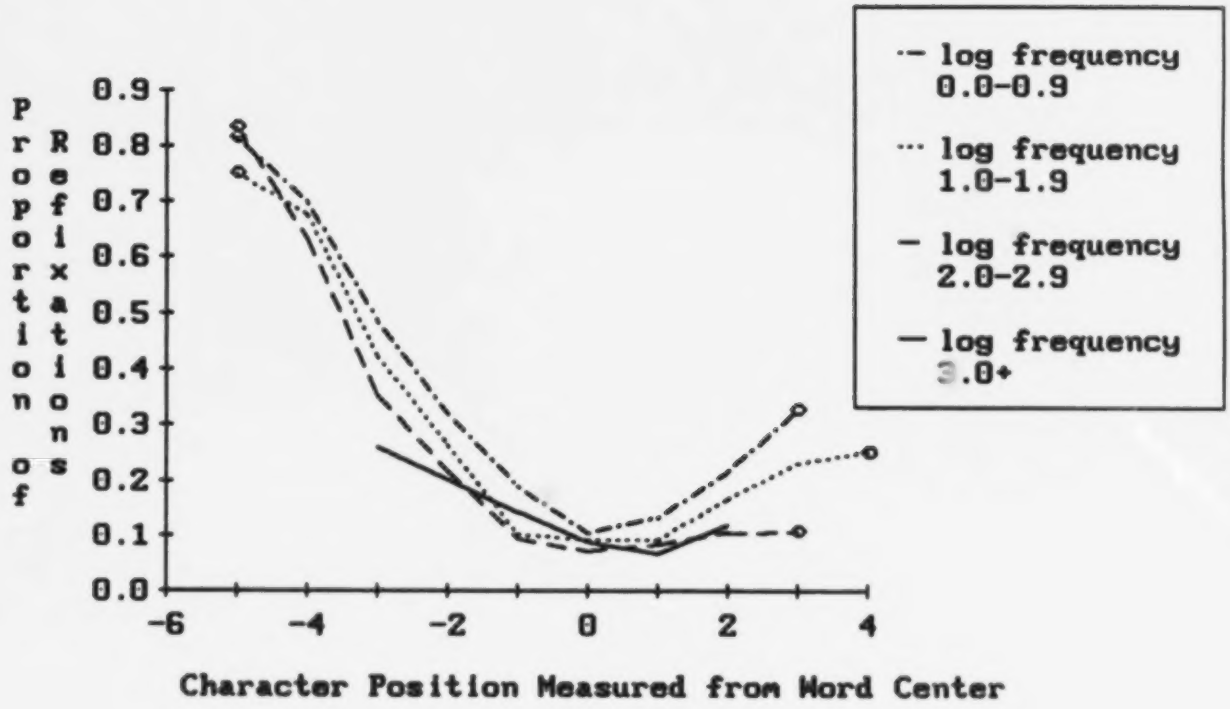
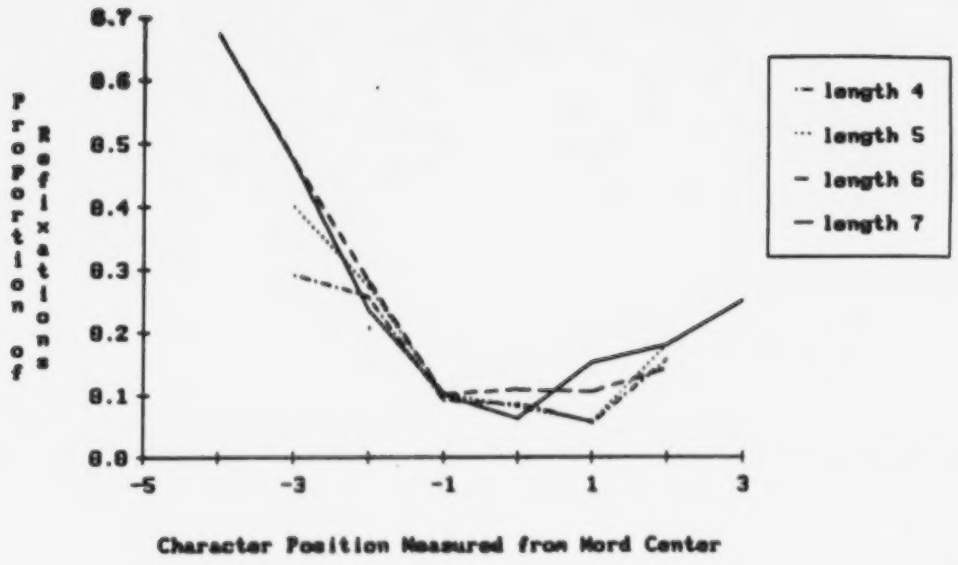

log frequency 0.0-0.9
log frequency 1.0-1.9
log frequency 2.0-2.9
log frequency 3.0+
Proportion of Refixations
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
-6
-4
-2
0
2
4
Character Position Measured from Word Center

Figure 5


0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
Proportion of Refixations
-5
-3
-1
1
3
Character Position Measured from Word Center
length 4
length 5
length 6
length 7

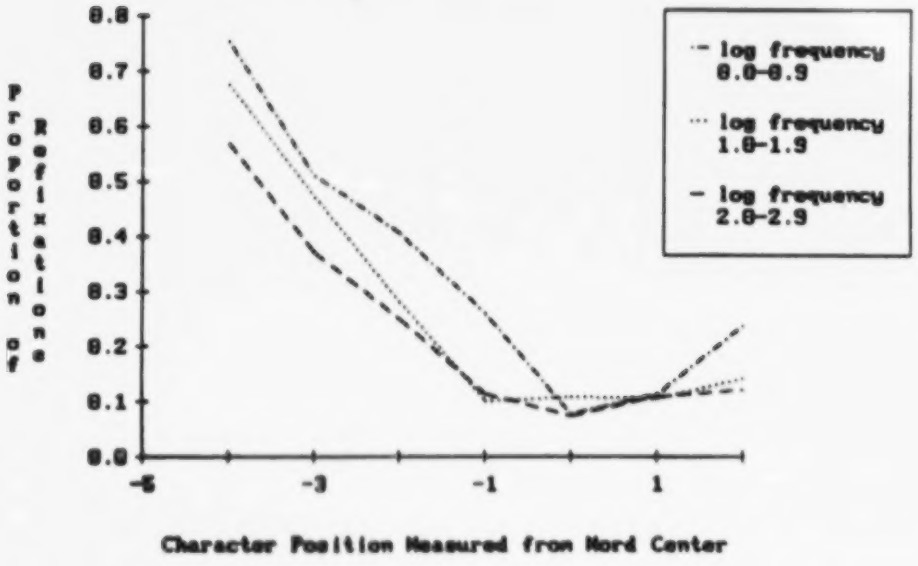

0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
Proportion of Refixations
-5
-3
-1
1
Character Position Measured from Word Center
log frequency 0.0-0.9
log frequency 1.0-1.9
log frequency 2.0-2.9

Figure 6